

## ARTICLE

Received 13 Jul 2014 | Accepted 8 Sep 2014 | Published 14 Oct 2014

DOI: 10.1038/ncomms6172

# Host-parasite network structure is associated with community-level immunogenetic diversity

Shai Pilosof<sup>1</sup>, Miguel A. Fortuna<sup>2</sup>, Jean-François Cosson<sup>3</sup>, Maxime Galan<sup>3</sup>, Chaisiri Kittipong<sup>4</sup>, Alexis Ribas<sup>5</sup>, Eran Segal<sup>6</sup>, Boris R. Krasnov<sup>1</sup>, Serge Morand<sup>7,8,9</sup> & Jordi Bascompte<sup>2</sup>

Genes of the major histocompatibility complex (MHC) encode proteins that recognize foreign antigens and are thus crucial for immune response. In a population of a single host species, parasite-mediated selection drives MHC allelic diversity. However, in a community-wide context, species interactions may modulate selection regimes because the prevalence of a given parasite in a given host may depend on its prevalence in other hosts. By combining network analysis with immunogenetics, we show that host species infected by similar parasites harbour similar alleles with similar frequencies. We further show, using a Bayesian approach, that the probability of mutual occurrence of a functional allele and a parasite in a given host individual is nonrandom and depends on other host-parasite interactions, driving co-evolution within subgroups of parasite species and functional alleles. Therefore, indirect effects among hosts and parasites can shape host MHC diversity, scaling it from the population to the community level.

<sup>1</sup>Mitrani Department of Desert Ecology and Albert Katz International School for Desert Studies, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion 8499000, Israel. <sup>2</sup>Integrative Ecology Group, Estación Biológica de Doñana (EBD-CSIC), Seville 41092, Spain. <sup>3</sup>INRA, UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro), Campus International de Baillarguet, CS 30016, 34988 Montpellier-sur-Lez cedex, France. <sup>4</sup>Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok 10400, Thailand. <sup>5</sup>Biodiversity Research Group, Faculty of Science, Udon Thani Rajabhat University, Udon Thani 41000, Thailand. <sup>6</sup>Department of Computer Science and Applied Mathematics, Weizmann Institute, Rehovot, 76100, Israel. <sup>7</sup>Centre National de la Recherche Scientifique—Institut des Sciences de l'Évolution, Université Montpellier 2, Montpellier 34095, France. <sup>8</sup>Centre de coopération International en Recherche Agronomique pour le Développement, Animal et Gestion Intégrée des Risques, Campus de Baillarguet, F-34093 Montpellier Cedex 5, France. <sup>9</sup>Centre d'Infectiologie Christophe Mérieux du Laos, PO Box 3888, Samsenthai Road, Vientiane, Lao PDR. Correspondence and requests for materials should be addressed to S.P. (email: shainova@gmail.com).

Species interactions are a source of selective pressures that can maintain genetic variability in host populations, parasitism being a classic example<sup>1–3</sup>. In particular, this is true for the major histocompatibility complex (MHC) in vertebrates. The MHC is a highly polymorphic region in the genome comprising closely linked genes. These genes encode cell-surface glycoproteins that recognize and present foreign antigens, such as those of invading parasites, to T lymphocytes thereby initiating an immune response<sup>4</sup>. Thus, at the host individual level, certain MHC alleles confer resistance/susceptibility to some parasites but not to others<sup>5–7</sup>.

Parasite-mediated selection drives MHC polymorphism at the host population level because of the evolutionary interplay between parasites and MHC-encoded proteins within a host individual<sup>8,9</sup>. Indeed, several studies have shown that in a host population, parasite diversity—the number and identity of parasite species and their abundance in a host population—is correlated with MHC diversity—the number of MHC alleles in a host and their identity—but not with genetic diversity assessed with neutral markers such as microsatellites or mitochondrial genes<sup>5,10</sup>.

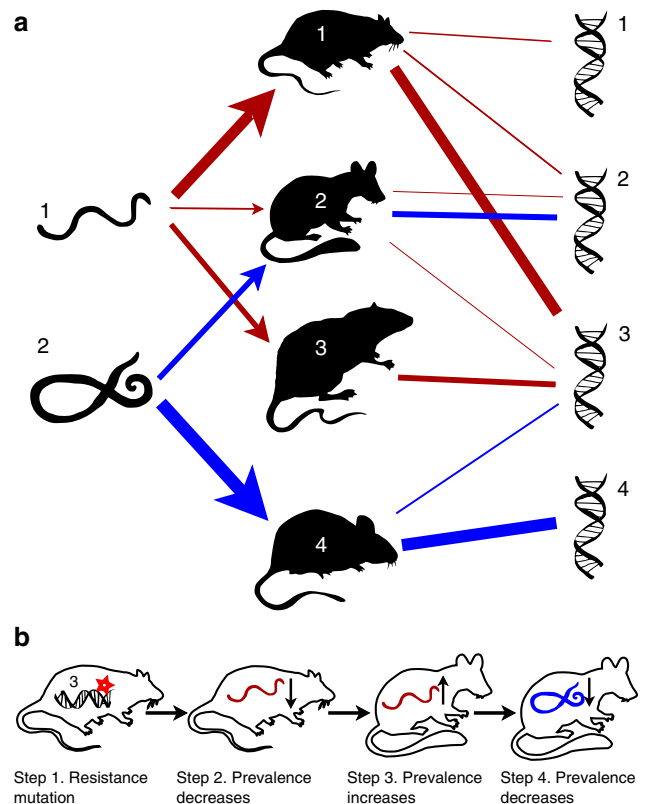
At the species level, studies of host interspecific variation in MHC are scarce but have confirmed a positive correlation between parasite species richness and MHC richness across species<sup>11,12</sup>. However, these studies were based on data collected from the literature, dealt with host species that were not simultaneously surveyed both for parasites and MHC alleles, and did not consider the functional diversity of alleles. More importantly, the way interactions among hosts and parasites are distributed was ignored and thus no study examined the association between multiple host and parasite interactions and MHC diversity. Although few studies have explored the effects of multiple parasites on MHC diversity<sup>7,10,13</sup>, they were conducted within a population of a single host.

Interactions between multiple hosts and parasites are often depicted as a host–parasite network. Considering network structure is crucial because, in nature, species do not live in isolation but rather belong to a community where species affect each other directly or indirectly<sup>14,15</sup>. In this community-wide context, the prevalence of a given parasite species in a given host species may depend on its prevalence in other host species in the community. Therefore, the selective pressure exerted by the parasite may be influenced by indirect species interactions and may itself influence diversity of MHC alleles directly involved in the host–parasite interaction (Fig. 1). That is, the structure of the host–parasite network may correlate with MHC diversity among host species. Using a unique data set on a community of rodents, their helminth parasites and their MHC alleles, we tackle here, for the first time, this largely unexplored research topic by integrating complex network theory with immunogenetics.

We show that host species infected by similar parasites harbour similar MHC functional alleles with similar frequencies after controlling for phylogenetic effects. Also, the indirect effects among hosts and parasites influence the probability of mutual occurrence of a functional MHC allele and a parasite in a given host individual. Therefore, the ecology of host–parasite interactions and the evolution of MHC are intertwined, and indirect effects among hosts and parasites can shape host MHC diversity, scaling it from the population to the community level.

## Results

**Data set.** An extensive helminth survey was conducted in six localities in southeast Asia (Supplementary Fig. 1) on 950 specimens of 11 murinae rodent species within the framework of the Community of Rodents and their Pathogens project



**Figure 1 | An example for the interplay between host-parasite and host-supertype networks.** (a) On the left, host–parasite interactions between helminths and rodent species (ecological network). Arrow width is proportional to the prevalence of a given parasite on a host. On the right, the immunogenetic host–supertype network in which interactions represent the frequencies (indicated by line width) of supertypes harboured by individuals of a given host that are infected by each parasite (indicated by different colours). (b) A theoretical example showing how in a community-wide context, species and MHC supertypes affect each other indirectly with possible cascading indirect effects. Imagine that a mutation for resistance (depicted by a red star) occurs in an allele from supertype no. 3 in the population of host no. 1 (step 1), causing the prevalence of parasite no. 1 to decrease in that host (step 2). Parasite no. 1 now switches to host no. 2 (step 3), thus reducing the prevalence of parasite no. 2 in host no. 2 due to competition (step 4).

(www.ceropath.org)<sup>16</sup>. Of these, 440 individuals were parasitized by 26 helminth taxa (22 identified to species level and 4 to family level) (Supplementary Tables 1–3).

Infection usually depends on species co-occurrence in space and time. Also, hosts of the same species but from different populations may be exposed to different selective pressures, thereby creating variation in MHC due to local adaptation. We therefore first analysed data using rodent individuals captured in a single locality (Loei). Despite geographical distance between the localities, they all belong to the same Indochina bioregion and are rather homogeneous in their environmental conditions<sup>17</sup>. Furthermore, all rodent species from other localities occurred in Loei (Supplementary Table 1). Therefore, we pooled data from the six localities and considered rodents as belonging to a single community to increase statistical power. This did not change our results qualitatively (see below) and we thus focus on the results of analyses done on pooled data.

Each individual rodent was genotyped using a portion of the MHC class II *DRB* exon-2 functional gene that recognizes extracellular parasites such as helminths<sup>18</sup>. Foreign antigens are recognized by particular amino acids of the MHC molecule called

the antigen-binding site. Some MHC alleles code for MHC molecules that bind an overlapping gamma of peptides. This happens when the antigen-binding sites of the coded molecules possess amino-acid sequences with similar physical and biochemical properties. Such alleles are thus functionally similar and can be classified into supertypes based on their antigen-binding site amino-acid sequences<sup>19–21</sup>. Because we were interested in functional diversity, and to avoid sparse networks caused by an excess of alleles with very few interactions, we first clustered functionally similar alleles into supertypes following refs 22,23 (see Methods section and Supplementary Fig. 2). This resulted in 24 supertypes but we omitted from statistical analysis one supertype (no. 23) that was not harboured by any infected individual, and thus 23 supertypes remained for analysis.

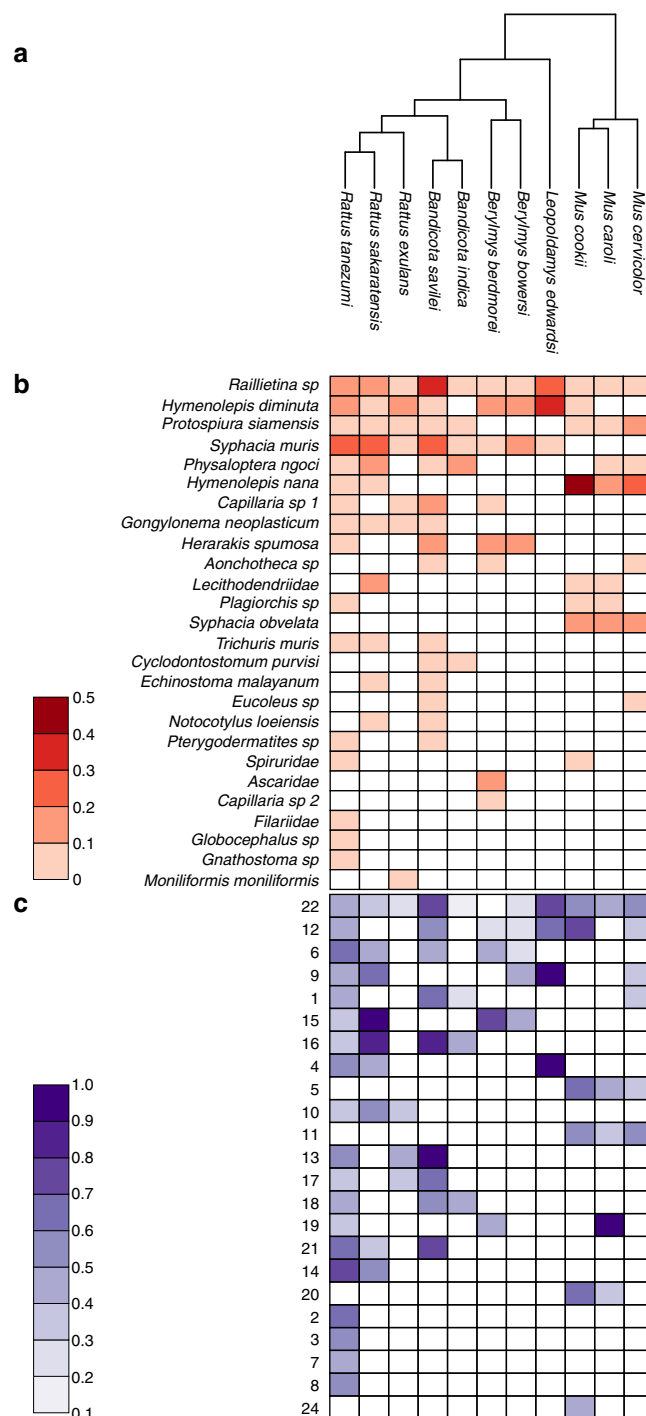
### Host-parasite and host-supertype networks are correlated.

We constructed a host-parasite ‘ecological network’ and a host-supertype ‘immunogenetic network’ in which the value of interactions were set as parasite prevalence or supertype frequency, respectively (matrix cell values in Fig. 2). If parasite-mediated selection is associated with the structure of the host-parasite network, then there should be a correlation between the identity and prevalence of parasite species and the identity and frequency of supertypes that hosts are infected by or harbour, respectively. Yet, since host species share a common evolutionary history, we have to account for the role that host phylogeny plays in the distribution of supertypes among host species (Fig. 2). We applied a partial Mantel’s test to quantitative dissimilarity matrices based on the ecological and immunogenetic networks, while using a matrix of pairwise phylogenetic distances to control for phylogenetic effects<sup>24</sup>, and found a positive and significant correlation between the host-parasite and host-supertype matrices ( $r=0.62$ ,  $P=0.001$ ; Supplementary Fig. 3). This means that two host species, infected by similar parasite species with a similar prevalence, tend to harbour the same MHC supertypes with similar frequencies. These results were robust to host abundances, to the number of parasites that infect a host, to taxonomic resolution of parasite species, to different phylogenetic trees and to statistical ‘group effects’ (Supplementary Methods).

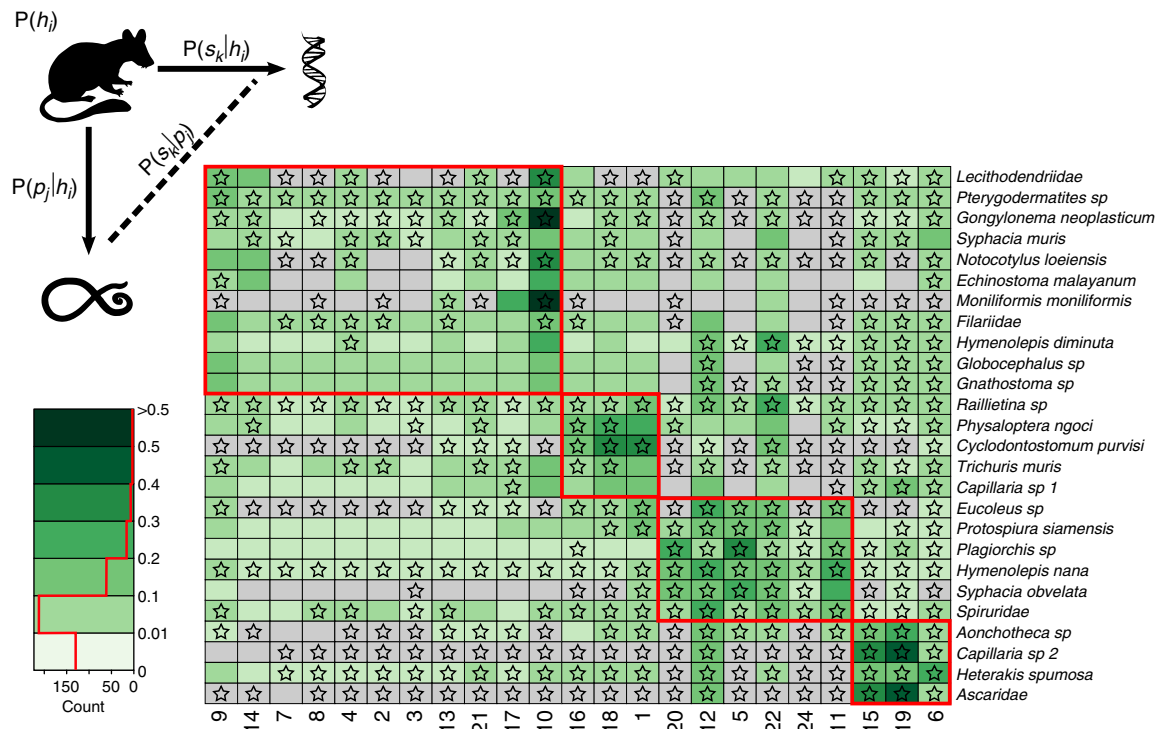
### Parasite-supertype associations across hosts are nonrandom.

We further explored this first evidence using a Bayesian approach that quantified parasite-supertype associations across host species (Fig. 3). This analysis estimates the probability of an individual host infected by a given parasite species to harbour a given supertype. This probability can be interpreted as a proxy for the resistance/susceptibility that a given supertype provides against a given parasite as compared with all other supertypes across hosts. A high probability means that in hosts where the parasite occurs, the supertype also occurs with high probability, and thus individuals harbouring that supertype are susceptible to the parasite. The advantage of this approach is threefold. First, we represent the same data in a different way, and obtaining similar results will strengthen our conclusions. Second, we can gain a new insight into our system: the probability of obtaining a given supertype  $s_k$  given a parasite  $p_j$ , calculated across hosts (see Fig. 3 and Methods section). Third, the calculation of the probability of a given parasite-supertype association is based on all host-parasite interactions, thus considering the structure of the host-parasite network.

The resulting parasite-supertype network (Fig. 3) contained all 23 supertypes and 26 parasites, with connectance (proportion of realized interactions out of all possible ones) of 72%, indicating that 28% of the potential supertype and helminths associations



**Figure 2 | Host phylogeny and the host-parasite and host-supertype matrices.** (a) Host phylogeny. An ultrametric tree is presented for convenience. (b) Parasite prevalence: the fraction of individuals of a given host species (columns) infected by a given parasite (rows). Rows are sorted from the most generalist parasite to the most specialist one. (c) MHC supertype frequency: number of infected individuals of a given host species (columns) with a given supertype (rows), divided by the total number of individuals of that host harbouring that supertype. One supertype (no. 23) was not considered as it was not harboured by any infected individuals. Rows are sorted from the most widely distributed supertype among hosts to the most species-specific.



**Figure 3 | Parasite-supertype association matrix.** Matrix values depict the probability of sampling an individual of any host species infected by parasite species  $p_j$  that harbours supertype  $s_k$ — $P(s_k|p_j)$ . This probability of association is a proxy for the resistance a given supertype provides against a given parasite regardless of the identity of the host species, with higher probabilities indicating lower resistance. This is indicated by the green shades in the legend (the darker the shade of green colour the lower the resistance). The histogram in the legend (red line) depicts the distribution of association values. Grey cells indicate that the supertype and parasite did not co-occur in any host. Stars indicate statistically significant differences of the observed parasite-supertype probabilistic associations from random when compared with 1,000 reshuffled matrices. Modules of superotypes and parasites indicate subgroups with higher potential to co-evolve and are marked with a red border. The matrix was calculated using a Bayesian network with three nodes: host species, parasite species and MHC supertype, by combining the following conditional probabilities: probability of sampling a parasite species  $p_j$  on a given host species  $h_i$ ,  $P(p_j|h_i)$ , the probability of sampling a supertype  $s_k$  on a host species  $h_i$ ,  $P(s_k|h_i)$  and the probability of obtaining a host  $P(h_i)$ .

were not realized in any host. The values of the probabilistic associations were not affected by host abundance (Supplementary Methods). The first important result evident from this analysis is that the association between superotypes and parasites is variable across the host community: Whereas the association of some parasites to superotypes is homogeneous for all superotypes, other parasites are strongly associated with specific superotypes. For instance, hosts are susceptible to *Syphacia muris* to a similar extent, regardless of the superotypes they harbour but hosts that harbour supertype no. 10 are more susceptible to *Moniliformis moniliformis* than others. Similarly, the resistance some superotypes (for example, superotypes no. 7 and no. 8) provide is homogeneous to all parasites, whereas others (for example, superotypes no. 10 and no. 19) do not confer resistance to specific parasites.

This analysis constitutes the first across-species analysis of the associations between multiple hosts, parasites and superotypes, and indicates that the diversity of MHC superotypes across hosts is associated with the structure of the host–parasite network. To further examine whether the observed parasite–supertype probabilistic associations are correlated with the structure of the host–parasite network, we compared the parasite–supertype matrix (Fig. 3) to 1,000 matrices recalculated after reshuffling the host–parasite interactions ( $\mathbf{M}_{\text{rand}}$ ). If the observed associations between parasites and superotypes ( $\mathbf{M}_{\text{obs}}$ ) are not correlated with the host–parasite network structure, then they are determined by a random process and should resemble the values obtained in the random matrices. We compared cell values of  $\mathbf{M}_{\text{obs}}$  with distributions of their respective cell values from  $\mathbf{M}_{\text{rand}}$  obtained by two null

models that reshuffle only the values or the values and structure of the host–parasite matrix. We found that 47% and 55% of the associations, respectively, significantly differed from random ( $P < 0.05$ ; two-tailed permutation tests), accounting for 66% of the parasite–supertype associations for the two null models combined (Fig. 3 and Supplementary Figs 4–6). These results were robust to host abundance and parasite taxonomic resolution (Supplementary Methods). This provided a second strong evidence for the interplay between host–parasite network structure and the diversity of MHC in different hosts.

**Networks are nested.** If indeed there is an interplay between the ecological and immunogenetic networks, then we should be able to predict a structural property of the immunogenetic host–supertype network based on the ecological host–parasite network. One common structural property of ecological networks is nestedness. In a nested host–parasite network, hosts infected by few parasite species (further referred to as ‘specialist’ hosts) tend to be parasitized by complete subsets of the parasites that infect more ‘generalist’ hosts—a common pattern in ecological networks<sup>25,26</sup>. The structure of the host–parasite network studied here was significantly nested (NODF = 51.8,  $P < 0.001$ ; ref. 27). Thus, a substantial proportion of interactions exists among a ‘core’ of generalist hosts and parasites, which are probably those most associated with the diversity of MHC. This finding leads us to predict that hosts with low MHC richness (that harbour few superotypes) will harbour subsets of the superotypes harboured by MHC-rich hosts, resulting in a nested host–supertype network. This prediction was confirmed (NODF = 50.3,  $P < 0.001$ ).



The nested structure is of particular interest because parasite species differ in their host specificity. Thus, parasites specific to one or few hosts may encounter only a limited spectrum of supertypes. Hence, we can predict a nested structure in the parasite–supertype network because, owing to the nested structure of the host–parasite network, parasites infecting subsets of the hosts that more generalist parasites infect should also encounter subsets of supertypes harboured by those hosts. This prediction was also confirmed (NODF = 78.7,  $P < 0.001$ ). However, nestedness is a qualitative measure that ignores parasite identity. We can further predict that if specificity is driving the observed patterns, then generalist parasites (infecting multiple hosts) will also be associated with many supertypes. If no such correlation is found, then parasite specificity is not a major factor to consider in our system. An additional analysis of parasite specificity that considered the quantitative nature of the networks and parasite identity showed that parasites specific to particular hosts are not necessarily specific to particular supertypes (Pearson's correlation;  $r = 0.2$ ,  $P = 0.34$ ; Supplementary Fig. 7). This happens because some parasites with high host specificity infect hosts with high MHC diversity (for example, *Gnathostoma* sp. is specific to *Rattus tanezumi* that harbours many supertypes; Fig. 2). Taken together, these two analyses of specificity suggest that the interplay between the ecological and immunogenetic networks is not driven by patterns of parasite specificity, but that there is a predictable nested pattern in the selection regime between parasites and supertypes, mediated through hosts.

**Parasite–supertype co-evolution.** From an evolutionary perspective, the molecular affinity between the proteins coded by given supertypes and parasites should lead to tight associations between subgroups of parasites and supertypes in our network. We examined this hypothesis using modularity analysis and found that the parasite–supertype network had four modules (Fig. 3). Supertypes and parasites within the same module interact with each other to a greater extent than with supertypes or parasites from other modules, suggesting a higher potential of parasite–supertype coevolution. The modularity of the observed parasite–supertype association network ( $Q = 0.37$ ) was significantly higher than that of random networks assembled as above using both null models ( $P < 0.001$  for both null models), indicating that the modular structure represented true co-evolutionary interactions rather than random associations between hosts and parasites.

**Analysis of a single community.** In Loei, there were 342 rodent individuals (abundance ranged between 8 and 84; Supplementary Table 1) belonging to 11 species. These species were parasitized by 18 of the 26 helminths taxa (70%) in the pooled data set.

There was a positive and significant correlation between the two host–parasite and host–supertype matrices, after controlling for phylogenetic effects ( $r = 0.7$ ,  $P = 0.001$ ). This correlation was maintained when controlling for host abundances and statistical ‘group effects’ (Supplementary Methods).

The parasite–supertype matrix contained all 18 helminths taxa and 17 supertypes. The connectance of this network was 65%. When compared with the two null models that reshuffle only the values or the values and structure of the host–parasite matrix, we found that 39% and 58% of the associations, respectively, significantly differed from random ( $P < 0.05$ ; two-tailed permutation tests), accounting for 67% of the parasite–supertype associations for the two null models combined.

The modularity of the observed parasite–supertype association network ( $Q = 0.37$ ) was significantly higher than that of random networks ( $N = 1,000$ ) assembled as beforehand using the two

null models that shuffle interaction values or values and network structure ( $P < 0.001$  for both). The network had three modules.

The host–parasite, host–supertype and parasite–supertype networks were all significantly nested (NODF = 48.4, 44.4 and 80.8,  $P < 0.01$  for all three networks). As with the pooled data, parasites specific to particular hosts were not specific to particular supertypes (Pearson's correlation;  $r = 0.37$ ,  $P = 0.13$ ).

## Discussion

This study is the first to consider the structure of the associations between multiple parasites and hosts in the context of functional MHC diversity, leading to novel insights regarding the ecology and evolution of MHC. Specifically, our results lead to the conclusion that considering the population level alone is not sufficient to understand parasite-mediated selection. Even the across-host positive correlation between MHC allele richness and parasite richness demonstrated in earlier studies<sup>11,12</sup> gives only a narrow picture of the complex evolutionary process. Instead, the ecological pattern in which host–parasite interactions are organized is crucial to understand adaptive evolution of MHC.

It is evident from our data that the direct and indirect effects hosts and parasites exert on each other are correlated with more than half of the variability in MHC diversity resulting from parasite-mediated selection. Yet, some parasite–supertype associations were not affected by the host–parasite network structure, indicating that associations can also depend solely on the parasite and supertype. This can happen, for example, when a parasite is specific to a host that has low MHC diversity resulting in independence from other host–parasite interactions.

We note that because the parasite–supertype association network is probabilistic in nature, devising a clear-cut criterion or threshold for what is a ‘biologically meaningful’ parasite–supertype association may not be the best practice, and should also greatly depend on the research question and system. In fact, because this is a result of two interconnected networks (Fig. 1), all associations are ‘meaningful’ to some extent because they all mutually affect each other. This reflects well the true biological complexity in interactions between parasites and supertypes mediated through hosts. Parasite–supertype association networks created for other communities may have other properties than those of the network described here, and this is a rich area for future research.

Nature is dominated by complex dynamics between species co-evolution and the structure of the networks they assemble<sup>28</sup>. Here, host–parasite co-evolution is also dependent on MHC. Hosts that evolve resistance remove interactions from the host–parasite network (or weaken existing ones) and parasites that evolve new pathways of infection add interactions to the host–parasite network (or strengthen existing ones) through genetic change in MHC alleles. When such dynamics occur at the community level, they can simultaneously affect multiple hosts, parasites, MHC alleles and the interactions among them. For example, an overall reduction in the genetic resistance in a host population to a parasite owing to changes in allele frequencies will cause an increase in the parasites' prevalence in the host population. If the parasite is not a strict host specialist, then selection imposed by the parasite on individuals of a second host species may also increase, thereby promoting MHC diversity in that host. This process can be extended to multiple hosts and parasites, eventually resulting in a complex network of indirect effects between hosts, parasites and MHC supertypes, whereby parasite-mediated selection on MHC supertypes of one or few hosts can have cascading effects on the MHC diversity of the whole host community.

Hence, and following previous evidence that genetic diversity can shape community structure<sup>29</sup>, we propose to extend current hypotheses explaining mechanisms of parasite-mediated selection<sup>9</sup> to include multiple host–parasite interactions. For instance, according to the ‘rare-allele advantage hypothesis’<sup>9</sup>, rare alleles in the host population confer better resistance than more frequent alleles because parasites have not had the time to evolve avoiding mechanisms (resistance). The rise of such an allele in a particular host species would weaken the host–parasite interaction in the network, with cascading effects on other species. With time, the rare allele becomes more frequent in the population and thus its selective advantage decreases as parasites evolve resistance. This would strengthen the host–parasite interaction, again affecting other species in the community. Once the allele becomes rare again, parasite adaptation for it also decreases, causing the selective advantage of the allele to increase again. Because changes to one interaction in the network have cascading effects on other interactions, there should be a dynamical interplay between the structural patterns of the host–parasite and host–supertype networks eventually affecting the co-evolution between parasites and MHC. This view is supported by the predictable nested pattern of parasite–supertype associations and by the modular network structure that indicates high co-evolutionary potential between particular parasites and supertypes.

The association between host–parasite network structure and the diversity of MHC supertypes suggests that a perturbation to the structure of the host–parasite network may result in a concomitant perturbation to the host–supertype network. Human-driven environmental changes can alter the interactions between hosts and parasites<sup>30</sup>. For example, rodents and their associated parasites can invade disturbed areas, possibly creating new interactions between local rodents and invading parasites<sup>31</sup>. Such alteration to the host–parasite network may have consequences to the selection pressure imposed by parasites and to the evolution of MHC supertypes through indirect cascading effects. Because rodents are carriers of zoonotic parasites and pathogens<sup>32,33</sup> (Supplementary Table 3), the ecological-evolutionary dynamics between altered host–parasite interactions and MHC may have implications for risk of zoonotic diseases<sup>34</sup>. Indeed, the rapid evolutionary potential of both hosts and parasites<sup>34</sup> makes genetic change possible even on ecological timescales.

This study is the first to integrate tools and theory of network ecology with immunogenetics, providing novel insights into the evolution of functional MHC alleles in the context of community ecology. In particular, our results lead us to conclude that more than half of the variability in MHC diversity may be shaped by host–parasite interactions in the community, with reciprocal effects of the changes in MHC diversity on host–parasite interactions. This scales our view of MHC diversity from the population level to the community level. The complexity inherent to this model of network-based parasite-mediated selection promotes MHC diversity through indirect pathways, and suggests that the structure of host–parasite interactions plays a major role in the evolution of host immunity at the community level.

## Methods

**Ethics statement.** All work with animals was approved by the French National Research Agency, project ANR 07 BDIV 012. Animals were treated in accordance with the guidelines of the American Society of Mammalogists and with the European Union legislation (Directive 86/609/EEC). None of the rodent species investigated were on the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) list or on the Red List of the International Union for Conservation of Nature (IUCN). Each trapping session was validated by the national, regional and local health authorities, and under an oral agreement of local land owners. Approval notices for trapping and investigating rodents were given by the Ethical Committee of Mahidol University, Bangkok, Thailand, number 0517.1116/661 and by the Ministry of Health Council of Medical Sciences, National Ethics Committee for Health Research Lao PDR, number 51/NECHR based on the

validation of the rodents trapping book protocols of CERoPath<sup>16</sup>. Cambodia has no ethics committee overseeing animal experimentation. In addition, regional approval was obtained from the regional Head of Veterinary Service (Hérault, France) for sampling and killing rodents and harvesting rodent tissues (approval no. B 34-169-1).

**MHC genotyping.** Since previous studies have shown an association between helminthic infections and certain MHC alleles<sup>5,7,35,36</sup>, each individual in the data set (Supplementary Data 1) was genotyped using a portion of the MHC class II *DRB* exon-2 functional gene that recognizes extracellular parasites such as helminths, by two independent runs of 454 GS-FLX sequencing following the procedure detailed in ref. 18. In brief, the locus was amplified by PCR using a modified version of the primers JS1 (5′-GCCTCCCTCGCGCCATCAG-3′) and JS2 (5′-GCCTTGCCAGCCCCGCTCAG-3′), which amplify a 171-bp fragment (excluding primers) of exon 2 from the *DRB* gene in several mammal species. These primers include in the 5′ position an individual-specific multiplex identifiers and the 454 adaptors required for the emulsion PCR and the 454 GS-FLX sequencing<sup>18</sup>. The SEquence Sorter & AMPlicon Explorer software (SESAME ver 1.1B; ref. 37) was used to sort the sequences (that is, individual assignment of the sequences and removal of artefactual variants generated during PCR, emulsion PCR and 454 sequencing), and to determine the genotype of each individual. Gene duplication is common in MHC genes, and, in theory, an individual can have more than two alleles. However, this was ruled out for all the 11 species studied<sup>18</sup>. Thus, individuals of the rodent species used here had a maximum of two alleles amplified by the PCR primers.

**Supertype clustering.** The validated MHC alleles were clustered with similar antigen binding into supertypes as proposed in ref. 22. The clustering to supertypes was based on 22 amino-acid sites detected as antigen-binding site on the basis of X-ray data of peptide class II protein complexes<sup>38,39</sup> and/or detected under positive selection using the HyPhy programme<sup>40</sup> included in the MEGA5 software<sup>41</sup>. Each of the 22 amino acids forming the antigen-binding site was characterized by five physicochemical descriptor variables: z1 (hydrophobicity), z2 (steric bulk), z3 (polarity), z4 and z5 (electronic effects).

Our final matrix for supertype classification had 276 rows representing unique alleles and 110 columns (5 physicochemical descriptors for 22 sites). We clustered the unique alleles into supertypes using a *K*-means clustering algorithm and model selection following the procedure described in ref. 23 with the function ‘find.clusters’ from package ‘adeget’ (version 1.3-9; ref. 42). In brief, the optimal number of supertypes (clusters) is defined as the minimal number of clusters after which the Bayesian Information Criteria (BIC) decreases by a negligible amount as indicated by the curve of BIC values as a function of cluster number (Supplementary Fig. 2). We used a manual classification criterion as recommended in the ‘adeget’ package reference. However, classification is a stochastic process (that is, there is no absolute true classification) and we thus validated our classification as follows. We ran 500 automated classifications using the option ‘smoothNgoesup’, in which the selected number of clusters is the *K* after which increasing the number of clusters leads to increasing the BIC in a loess-smoothed curve. Our manual classification to 24 supertypes was almost identical to the mean of the automated classification process and fell within its 95% confidence intervals (CIs; mean  $\pm$  s.d. 23.23  $\pm$  3.7; 95% CI (16,29)). Furthermore, we used a measure of normalized mutual information<sup>43</sup>, based on information theory, to examine the differences between the automated and manual classifications in terms of allele identities. The index quantifies the amount of certainty that we can have in our manual classification by observing an automated one. If the manual classification is identical to the automated one (clusters in the manual and automated classifications contain the exact same alleles), the index takes its maximum value of 1. If the classifications are totally independent (complete mismatch in clustering), its value is 0 (see details in ref. 43). The average mutual information across the 500 automatic classifications was 0.85, indicating that our classification was accurate.

In Loei, a classification of the 144 alleles found across individuals revealed 18 supertypes, one of which was not harboured by infected individuals and was omitted from analyses. Our manual classification to supertypes was almost identical to the mean of the automated classification and fell within its 95% CIs (mean  $\pm$  s.d. 18.28  $\pm$  2.4; 95% CI (12,21)). The average normalized mutual information across the 500 automatic classifications was 0.88, indicating that our classification was accurate. Furthermore, a comparison of supertype classification between Loei and the pooled data set resulted in a normalized mutual information of 0.84. This essentially indicated that by knowing the classification in the Loei data set, we can know the classification in the pooled data set and vice versa with 84% certainty, indicating a great overlap in the classification of alleles to particular supertypes even between the data sets.

We matched the two alleles of an individual with their corresponding supertypes such that each individual in the data set was also characterized by a pair of supertypes. To simplify, in this study, the terms genotype, homozygote and heterozygote refer to supertypes, not alleles, and are defined as follows. A supertype is a functional group to which an allele belongs, and a genotype is a combination of two supertypes (in case of a heterozygote) or one supertype (in case of a homozygote) to which the two alleles of an individual belong. For example, if alleles A and B were classified as supertype no. 1, and allele C was classified as supertype

no. 2, an individual bearing alleles A and B would be a homozygote with a genotype 1-1, whereas an individual bearing alleles B and C would be a heterozygote with a genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the parasite in question<sup>13</sup>. It is common in studies of MHC diversity to use only alleles/supertypes, rather than genotypes in statistical analyses<sup>5,6,44</sup>. However, owing to the lack of knowledge of the genetic mechanism, and because working with genotypes would create sparse networks, we used supertypes and validated our analyses with genotypes<sup>13</sup>. The results were qualitatively similar and we thus present only analyses done with supertypes. Analyses done using genotypes can be found in Supplementary Methods.

**Phylogeny.** We built a phylogenetic tree based on molecular data of the cytochrome *b* mitochondrial gene. We compiled cytochrome *b* sequences from the NCBI gene bank and used a maximum likelihood analysis with the GTR + G + I substitution model of molecular evolution with the aid of the function 'phymtst' in the R package 'ape'<sup>45</sup>. To ensure that our results were not affected by the way we constructed the tree, we re-ran our analyses with two other trees from refs 46,47, which did not change our results qualitatively (Supplementary Methods).

**Similarity analysis.** In the host–parasite 'ecological' network, we connected between a rodent species and a helminth when at least one rodent individual was infected by this helminth. Interaction weights (matrix cell values in Fig. 2a) were set as parasite prevalence—the fraction of individuals of a given rodent species infected by a given helminth. We described the diversity of supertypes across hosts with an 'immunogenetic' network (Fig. 1). We connected a supertype and a host if at least one infected individual of the host harboured the supertype. The weights of network interactions (matrix cell values in Fig. 2b) were determined as supertype frequency—the number of infected individuals of a given rodent with a given supertype, divided by the total number of individuals of that host harbouring the supertype. In calculating interaction weights, heterozygote individuals were counted twice (once for each of their supertypes), whereas homozygote individuals were counted once. This assured that interaction weights represented supertype frequencies rather than individual frequencies.

We examined whether hosts infected by similar parasites with a similar prevalence harbour similar MHC supertypes with similar frequencies, while controlling for phylogenetic effects, using a partial Mantel's test following ref. 24. For this, we created dissimilarity matrices for the host–parasite and host–supertype networks (matrices) using the Ružička index (RI), implemented in the R package 'vegan'<sup>48</sup>, defined as:  $RI = 2D/(1 + D)$ , where  $D$  is dissimilarity between hosts  $j$  and  $k$  given by  $D_{jk} = \sum[(x_{mj} - x_{mk})/\sum(x_{mj} + x_{mk})]$ , within which  $x_m$  is the prevalence of parasite  $m$  (or frequency of supertype  $m$ ) on hosts  $j$  and  $k$ , summed for all parasites infecting (or supertypes harboured by)  $j$  or  $k$ . The RI is a quantitative version of the Jaccard index and allows for direct comparisons between hosts<sup>49,50</sup>. An RI value of 1 indicates maximum dissimilarity, where no parasites (or supertypes) are shared between a pair of hosts, whereas an RI value of 0 indicates that parasites infect (or supertypes are harboured by) hosts with a similar prevalence (frequency).

To disentangle the effect of neutral species traits (or genes), which are assumed to be described by phylogeny, from that of parasite-mediated selection on MHC diversity, we created a phylogenetic distance matrix between pairs of host species using the function 'cophenetic.phylo' (R package 'ape', version 3.0-7), which calculates the sum of the length of the branches between pairs of species. We used this matrix as a control matrix in the partial Mantel's test, which examines the null hypothesis of no relationship between parasite diversity and MHC diversity using dissimilarity matrices. If host–parasite network structure is correlated with diversity of MHC supertypes across hosts (host–supertype network structure), we expect a positive and significant correlation between the host–parasite and host–supertype dissimilarity matrices (Supplementary Fig. 3).

**Bayesian analysis.** We calculated a matrix of parasite–supertype probabilistic associations,  $M_{obs}$ , by representing our system as a simple Bayesian network with three random variables (nodes): host, parasite and supertype (Fig. 3). For this Bayesian network, we recalculated the host–parasite and host–supertype matrices to create two conditional probability tables (rows sum to 1) representing the probability of sampling a parasite  $p_j$  or a supertype  $s_k$  given a host  $h_i$ . That is, the number of individuals of  $h_i$  infected with  $p_j$  divided by the total number of infected individuals of  $h_i$  (ecological matrix) and the number of infected individuals of  $h_i$  harbouring  $s_k$  divided by the total number of infected individuals of  $h_i$  (immunogenetic matrix). Biologically, the matrices represent the strength of infection of a parasite  $p_j$  on host  $h_i$  in comparison with all other parasites and the level of susceptibility/resistance conferred by a supertype  $s_k$  for a given host  $h_i$  in comparison with all other supertypes. The analysis also requires a conditional probability table for the host node of the Bayesian network. Because we did not have any preference for selecting a particular host, we assumed a uniform distribution for the probability of obtaining a host,  $P(h_i) = 1/11$  (total of 11 hosts). Given the Bayesian network and the conditional probability tables, we used the chain rule of Bayesian networks and Bayes theorem (see Proof 1 and equation (4)) to calculate the values of the matrix  $M_{obs}$ , depicting the probability of obtaining a supertype  $s_k$  given a parasite  $p_j$  across hosts (Fig. 3).

#### Proof 1:

$$P(s = s_k | p = p_j) = \quad (1)$$

$$\sum_{i \in H} P(s_k | h_i) \cdot P(h_i | p_j) = \quad (2)$$

$$\sum_{i \in H} P(s_k | h_i) \cdot \frac{P(p_j | h_i) \cdot P(h_i)}{P(p_j)} = \quad (3)$$

$$\sum_{i \in H} P(s_k | h_i) \cdot \frac{P(p_j | h_i) \cdot P(h_i)}{\sum_{i \in H} P(h_i) \cdot P(p_j | h_i)} \quad (4)$$

To examine whether the parasite–supertype associations (cell values of  $M_{obs}$ ) are correlated with the structure of the host–parasite network, we compared them with those derived from 1,000 matrices recalculated after randomizing the ecological matrix using two null models. The first null model, based on the IA algorithm from ref. 51, reshuffles only the values of the host–parasite network. The algorithm reassigns individuals randomly to matrix cells with probabilities proportional to observed row and column abundance totals until the matrix-wide total number of individuals is reached. The structure is not altered (that is, zero cells are kept the same in the randomized matrices). Therefore, non-observed host–parasite interactions are regarded to as 'forbidden'. The second algorithm reshuffles both the values and structure of the host–parasite matrix and is a modified version of the IF algorithm from ref. 51. This algorithm first shuffles the matrix according to the swap algorithm for binary matrices<sup>26</sup> and then assigns individuals randomly to matrix cells as in the previously described algorithm. In this case, non-observed host–parasite interactions are not regarded to as forbidden and are allowed to occur in the random matrices. We determined statistical significance (at the  $\alpha = 0.05$  level) as the proportion of association probabilities obtained by random that were either equal or larger or equal or lower than the observed association probabilities. Forbidden interactions were excluded from this calculation (Supplementary Fig. 5). We used a two-tailed permutation test because random probabilities of parasite–supertype associations could be either higher or lower than the observed probability. When calculating the proportion of significant interactions in the model that shuffles only interaction values, we omitted 60 parasite–supertype associations (11% of the total 520) that involved parasites that infect a single host species because the distribution of parasite–supertype associations resulting from the null model for such parasites consisted of a single value, always resulting in a significant parasite–supertype association. We note that both our null models were very conservative in the sense that they present elevated chances of not finding a significant pattern where it really does exist, as compared with models with less strict rules for network randomization<sup>26</sup>.

**Nestedness and modularity.** We calculated nestedness using the NODF index<sup>27</sup> implemented in the software Aninhando<sup>52</sup>. We calculated modularity using the QuanBiMod algorithm for quantitative bipartite networks<sup>53</sup> as implemented in the package 'bipartite' (version 2.03; ref. 54) in R using default values. To examine whether modules of supertypes and parasites were indeed the result of an evolutionary process rather than a random one, we explored whether our network was significantly more modular than random networks obtained by using both null models as described above (1,000 random networks per null model) using a one-tailed permutation test because random networks are expected to have lower modularity than the observed network.

**Parasite specificity.** We used the Paired Differences Index<sup>55</sup> as a quantitative measure of parasite specificity for hosts (in the host–parasite network) or supertypes (in the parasite–supertype network). The index contrasts the strongest interaction of a parasite on a host (or the strongest association of a parasite with a supertype), with the remaining interactions with other hosts (or remaining associations with other supertypes). We used the normalized version of the index, which divides the index for each species by the maximum of interactions. Thereby, parasite species can be compared among each other and values range between 0 (perfect generalist) and 1 (perfect specialist). We calculated the index using the function 'PDI' in the R package 'bipartite'.

**Data and R code.** The data and R code used in the analyses and to construct Figs 2 and 3 and Supplementary Tables and Figures are available as Supplementary Methods and Supplementary Data 2–9. The multiplexed raw sequences (.fasta files and .qual files) and the validated sequences of the 276 alleles are available from the Dryad Digital Repository: <http://doi.org/10.5061/dryad.mq7n0/1>.

#### References

- Reznick, D. N. A critical look at reciprocity in ecology and evolution: introduction to the symposium. *Am. Nat.* **181**, S1–S8 (2013).
- Thompson, J. N. in *Relentless Evolution* 121–143 (The University of Chicago Press, 2013).
- Carroll, S. P., Hendry, A. P., Reznick, D. N. & Fox, C. W. Evolution on ecological time-scales. *Funct. Ecol.* **21**, 387–393 (2007).



4. Janeway, C., Travers, P., Walport, M. & Shlomchik, M. *Immunobiology: the Immune System in Health and Disease* (Garland Science, 2005).
5. Schwensow, N., Fietz, J., Dausmann, K. H. & Sommer, S. Neutral versus adaptive genetic variation in parasite resistance: importance of major histocompatibility complex supertypes in a free-ranging primate. *Heredity* **99**, 265–277 (2007).
6. Srithayakumar, V., Castillo, S., Rosatte, R. C. & Kyle, C. J. MHC class II DRB diversity in raccoons (*Procyon lotor*) reveals associations with raccoon rabies virus (Lyssavirus). *Immunogenetics* **63**, 103–113 (2011).
7. Froeschke, G. & Sommer, S. Insights into the complex associations between MHC class II DRB polymorphism and multiple gastrointestinal parasite infestations in the striped mouse. *PLoS ONE* **7**, e31820 (2012).
8. Sommer, S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front. Zool.* **2**, 16 (2005).
9. Spurgin, L. G. & Richardson, D. S. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc. R. Soc. London B Biol. Sci.* **277**, 979–988 (2010).
10. Wegner, K. M., Reusch, T. B. H. & Kalbe, M. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J. Evol. Biol.* **16**, 224–232 (2003).
11. Garamszegi, L. Z. & Nunn, C. L. Parasite-mediated evolution of the functional part of the MHC in primates. *J. Evol. Biol.* **24**, 184–195 (2011).
12. Göüy de Bellocq, J., Charbonnel, N. & Morand, S. Coevolutionary relationship between helminth diversity and MHC class II polymorphism in rodents. *J. Evol. Biol.* **21**, 1144–1150 (2008).
13. Oliver, M. K., Telfer, S. & Pierny, S. B. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (*Arvicola terrestris*). *Proc. R. Soc. London B Biol. Sci.* **276**, 1119–1128 (2009).
14. Bascompte, J. Disentangling the web of life. *Science* **325**, 416–419 (2009).
15. Pilosof, S., Fortuna, M. A., Vinarski, M. V., Korallo-Vinarskaya, N. P. & Krasnov, B. R. Temporal dynamics of direct reciprocal and indirect effects in a host-parasite network. *J. Anim. Ecol.* **82**, 987–996 (2013).
16. Herbreteau, V. *et al. Protocols For Field And Laboratory Rodent Studies* (Kasetsart University Press, 2011).
17. Woodruff, D. S. Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. *Biodivers. Conserv.* **19**, 919–941 (2010).
18. Galan, M., Guivier, E., Caraux, G., Charbonnel, N. & Cosson, J.-F. A 454 multiplex sequencing method for rapid and reliable genotyping of highly polymorphic genes in large-scale studies. *BMC Genomics* **11**, 296 (2010).
19. Trachtenberg, E. *et al.* Advantage of rare HLA supertype in HIV disease progression. *Nat. Med.* **9**, 928–935 (2003).
20. Sette, A. & Sidney, J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics* **50**, 201–212 (1999).
21. Southwood, S. *et al.* Several common HLA-DR types share largely overlapping peptide binding repertoires. *J. Immunol.* **160**, 3363–3373 (1998).
22. Doytchinova, I. A. & Flower, D. R. *In silico* identification of supertypes for class II MHCs. *J. Immunol.* **174**, 7085–7095 (2005).
23. Sepil, I., Moghadam, H. K., Huchard, E. & Sheldon, B. C. Characterization and 454 pyrosequencing of major histocompatibility complex class I genes in the great tit reveal complexity in a passerine system. *BMC Evol. Biol.* **12**, 68 (2012).
24. Rezende, E. L., Lavabre, J. E., Guimarães, P. R., Jordano, P. & Bascompte, J. Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature* **448**, 925–928 (2007).
25. Bascompte, J., Jordano, P., Melián, C. J. & Olesen, J. M. The nested assembly of plant-animal mutualistic networks. *Proc. Natl Acad. Sci. USA* **100**, 9383–9387 (2003).
26. Fortuna, M. A. *et al.* Nestedness versus modularity in ecological networks: two sides of the same coin? *J. Anim. Ecol.* **79**, 811–817 (2010).
27. Almeida-Neto, M., Guimarães, P. R., Loyola, R. D. & Ulrich, W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**, 1227–1239 (2008).
28. Thompson, J. N. The coevolving web of life. *Am. Nat.* **173**, 125–140 (2009).
29. Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N. & Vellend, M. Ecological consequences of genetic diversity. *Ecol. Lett.* **11**, 609–623 (2008).
30. Laliberté, E. & Tylianakis, J. M. Deforestation homogenizes tropical parasitoid-host networks. *Ecology* **91**, 1740–1747 (2010).
31. Britton, J. R. Introduced parasites in food webs: new species, shifting structures? *Trends Ecol. Evol.* **28**, 93–99 (2013).
32. Meerburg, B. G., Singleton, G. R. & Kijlstra, A. Rodent-borne diseases and their risks for public health. *Crit. Rev. Microbiol.* **35**, 221–270 (2009).
33. Bordes, F. *et al.* The diversity of microparasites of rodents: a comparative analysis that helps in identifying rodent-borne rich habitats in Southeast Asia. *Infect. Ecol. Epidemiol.* **3**, 1–10 (2013).
34. Altizer, S., Harvell, D. & Friedle, E. Rapid evolutionary dynamics and disease threats to biodiversity. *Trends Ecol. Evol.* **18**, 589–596 (2003).
35. Harf, R. & Sommer, S. Association between major histocompatibility complex class II DRB alleles and parasite load in the hairy-footed gerbil, *Gerbillurus paeba*, in the southern Kalahari. *Mol. Ecol.* **14**, 85–91 (2005).
36. Lenz, T. L., Wells, K., Pfeiffer, M. & Sommer, S. Diverse MHC IIB allele repertoire increases parasite resistance and body condition in the Long-tailed giant rat (*Leopoldamys sabanus*). *BMC Evol. Biol.* **9**, 269 (2009).
37. Megléc, E. *et al.* SESAME (SEquence Sorter & Amplicon Explorer): genotyping based on high-throughput multiplex amplicon sequencing. *Bioinformatics* **27**, 277–278 (2011).
38. Robinson, J. *et al.* IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res.* **31**, 311–314 (2003).
39. Brown, J. H. *et al.* Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33–39 (1993).
40. Pond, S. L. K., Frost, S. D. W. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
41. Tamura, K. *et al.* MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
42. Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
43. Danon, L., Diaz-Guilera, A., Duch, J. & Arenas, A. Comparing community structure identification. *J. Stat. Mech. Theory Exp.* **9**, P09008 (2005).
44. Schad, J., Dechmann, D. K. N., Voigt, C. C. & Sommer, S. Evidence for the 'good genes' model: Association of MHC Class II DRB alleles with ectoparasitism and reproductive state in the neotropical lesser bulldog bat, *Noctilio albigentris*. *PLoS ONE* **7**, e37101 (2012).
45. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
46. Fabre, P., Hautier, L., Dimitrov, D. & Douzery, E. A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evol. Biol.* **12**, 88 (2012).
47. Pagès, M. *et al.* Revisiting the taxonomy of the Rattini tribe: a phylogeny-based delimitation of species boundaries. *BMC Evol. Biol.* **10**, 184 (2010).
48. Oksanen, J. *et al.* vegan: Community Ecology Package. <http://CRAN.R-project.org/package=vegan> (2012).
49. Speed, J. D. M., Cooper, E. J., Jónsdóttir, I. S., Van Der Wal, R. & Woodin, S. J. Plant community properties predict vegetation resilience to herbivore disturbance in the Arctic. *J. Ecol.* **98**, 1002–1013 (2010).
50. Tamás, J., Podani, J. & Csontos, P. An extension of presence/absence coefficients to abundance data: a new look at absence. *J. Veg. Sci.* **12**, 401–410 (2001).
51. Ulrich, W. & Gotelli, N. J. Null model analysis of species associations using abundance data. *Ecology* **91**, 3384–3397 (2010).
52. Guimarães, Jr P. & Guimarães, P. R. Improving the analyses of nestedness for large sets of matrices. *Environ. Model. Softw.* **21**, 1512–1513 (2006).
53. Dormann, C. F. & Strauss, R. A. A method for detecting modules in quantitative bipartite networks. *Methods Ecol. Evol.* **5**, 90–98 (2014).
54. Dormann, C. F., Gruber, B. & Fründ, J. Introducing the bipartite Package: Analysing Ecological Networks. *R News* **8**, 8–11 (2008).
55. Poisot, T., Canard, E., Mouquet, N. & Hochberg, M. E. A comparative study of ecological specialization estimators. *Methods Ecol. Evol.* **3**, 537–544 (2012).

## Acknowledgements

We thank Eli Geffen, Gili Greenbaum and Shai Kinast for their helpful discussions. We thank Yannick Chaval for help in field work, Sylvain Piry for the management and support of SESAME software, Julien Claude for help with creating phylogenetic trees and Irem Sepil for advice on supertype clustering. This study was supported by the French ANR, grants ANR 07 BDIV 012 (CERoPath project), ANR 11 CPEL 002 (BiodivHealthSEA project) and the programme APEGE-CNRS (NcRNARat project BFC 78167) and by the Institut National de la Recherche Agronomique (Projet innovant du département EFPA 2008). S.P. was funded by a PhD Fellowship from the Kreitman Foundation and a Aharon and Efraim Katzir Fellowship from the Israel Academy of Science and Humanities. M.A.F. was funded by a postdoctoral fellowship (JAE-Doc) from the Program 'Junta para la Ampliación de Estudios' co-funded by the Fondo Social Europeo (FSE). J.B. was supported by a European Research Council Advanced Grant. This is publication number 848 of the Mitrani Department of Desert Ecology.

## Author contributions

Project leadership: S.P., M.A.F., J.B.; MHC genotyping: J.-F.C. and M.G.; helminth identification: C.K. and A.R.; data collection management: S.M.; statistical analysis: S.P., M.A.F. and E.S.; interpretation of main findings: S.P., M.A.F., B.R.K., S.M. and J.B.; primary drafting of manuscript: S.P.; production of the final manuscript: S.P., M.A.F., B.R.K., S.M., J.-F.C., M.G. and J.B.



**Additional information**

**Accession codes:** The demultiplexed raw sequences generated in this study have been deposited in NCBI Short Read Archive (SRA) under the accession code SRP047116.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Pilosof, S. *et al.* Host–parasite network structure is associated with community-level immunogenetic diversity. *Nat. Commun.* 5:5172 doi: 10.1038/ncomms6172 (2014).